

DETECCIÓN DE PALLETS MEDIANTE TÉCNICAS DE VISIÓN POR COMPUTADOR

David Ramos Porres
davidramosburgos@gmail.com

J. Enrique Sierra
Universidad de Burgos, ASTI Mobile Robotics, jesierra@ubu.es

Resumen

En la actualidad los AGV's (Vehículos de Guiado Automático) conocen la posición de los pallets ya que almacenan sus ubicaciones. En caso de que se modifique la posición de algún pallet a través de algún elemento ajeno al sistema de control, por ejemplo sistemas de transporte manuales, la posición memorizada no coincidiría con la posición real del pallet, y el AGV no podría recogerlo. Por esta razón surge la necesidad de incorporar sensores ópticos en los AGV que sean capaces de identificar las ubicaciones de los pallets en tiempo real y así ajustar sus trayectorias.

En este trabajo se realiza un estudio del estado del arte sobre detección de pallets, se identifica el sistema de visión más adecuado para la aplicación, y se propone un nuevo algoritmo basado en técnicas de visión artificial. Los resultados experimentales contrastan la validez de la propuesta.

Palabras clave: AGV, pallet, detección, algoritmo

1 INTRODUCCIÓN

El acrónimo AGV proviene del término Vehículo de Guiado Automático. Estos son vehículos de transporte sin conductor empleados en la industria para automatizar los flujos intralógicos y las cadenas de producción, dotando a estos procesos de una gran flexibilidad espacial y temporal [7]. En este ámbito es importante que sea lo más autónomo posible. Para ello, en este estudio se busca que el AGV sea capaz de detectar la posible desviación del pallet con respecto a su posición esperada para que sea capaz de replanificar su trayectoria y coger el pallet sin dificultades.

En este trabajo se propone un algoritmo basado en la aplicación de diferentes técnicas de visión computacional: transformada de Hough[6], Canny, dilatación, erosión, etc [1][5]. Se comprueba la validez de la técnica mediante datos de un pallet real extraídos mediante un sensor Kinect v1.

La estructura de nuestra contribución es la siguiente. En la sección 2 se describe el estado del arte en esta temática. El método propuesto es descrito en la sección 3. Después los resultados son comentados en la sección 4. El documento finaliza con las conclusiones y líneas de trabajo futuras.

2 ESTUDIO DEL ESTADO DEL ARTE

Otros trabajos anteriores estudian el problema de la detección de pallets, a continuación se comentan algunos de los más relevantes.

El estudio [4] propone la instalación de una cámara 2D a la altura siempre del pallet. El objetivo es detectar el pallet tratando de detectar las esquinas de la cara frontal del pallet y saber diferenciar cuál es el taco central del pallet para dirigir el AGV hacia él. El algoritmo debe de ser capaz de detectar el pallet aun viéndolo parcialmente.

En la figura 1 se muestra un esquema a modo resumen del funcionamiento del algoritmo básico.

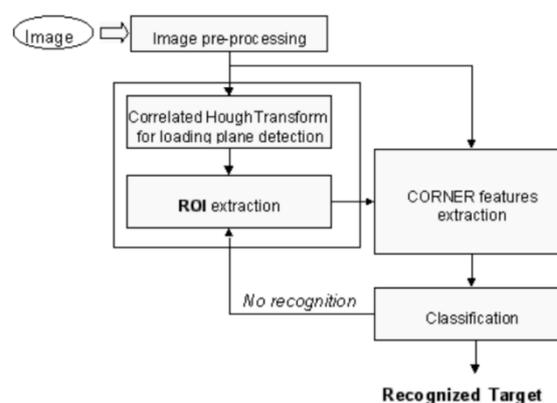


Figura 1: Resumen de la arquitectura del algoritmo

Uno de los estudios que más se puede aproximar a nuestro caso lo realiza NIST (National Institute of Standards and Technology) a través de un acuerdo cooperativo de investigación y desarrollo con Transrobotics un fabricante de AGVs [3].

Estos AGVs, destinados al llenado automático de camiones, llevan dos sensores láser: uno en la base del AGV en la parte inferior para evitar obstáculos y otro en la parte superior del AGV en la zona frontal. Este último lleva un motor incorporado para poder girar el láser 90° y así realizando un barrido poder obtener una nube de puntos 3D.

El algoritmo perseguía los siguientes objetivos: localizar el pallet y su orientación, verificar su correcto posicionamiento dentro del camión; y por último verificar que la carga estuviera bien posicionada en el pallet para evitar posibles caídas.

El algoritmo que utilizaron se puede resumir de la siguiente manera:

- Convertir la nube de puntos en el plano X e Y.
- El sistema predice la localización de las 4 esquinas del camión.
- Aplicar la transformada de Hough sobre el conjunto de puntos proyectados en 2D.
- Buscar las líneas más cercanas en el espacio de Hough que coincidan con la línea de la pared izquierda, derecha y posterior.
- Si cualquier pared del camión coincide con la predicción, se calcula la distancia desde el centro del sensor (0,0) a la pared.
- Para los puntos que están delimitados por las tres paredes del camión se calculan su distancia a las tres paredes.
- Se agrupa todos los puntos que tengan una distancia menor a la pared izquierda en comparación con la distancia a la pared derecha, estos puntos se consideran el taco izquierdo del pallet. Lo mismo se realiza para el taco derecho.

En la figura 2 se puede apreciar desde arriba la nube de puntos proyectada sobre un plano 2D del interior de un camión en el que hay un pallet al fondo a la izquierda:



Figura 2: Nube de puntos del interior del camión con un pallet al fondo a la izquierda

En[2] se propone el uso de un láser y el de una cámara a color correctamente calibradas entre sí. Los dos sistemas están montados de forma próxima.

Para detectar el pallet de forma aproximada se buscan los tres segmentos frontales del pallet, para ello se

compara con una imagen plantilla. Antes de compararlo con la plantilla se debe calcular el ángulo del pallet para poder rotar la plantilla y así compararlo.



Figura 3: a) Plantilla b) c) d) Pallets con diferentes ángulos

Estas orientaciones se calculan con la transformada de Hough. Hay que destacar que solo se cogen rectas que al menos contengan dos segmentos. Si detectamos varias líneas con diferentes pendientes, se crea un modelo por cada línea detectada y se selecciona el modelo que mejor se ajuste, figura 4.

Se inicializa la búsqueda con la imagen inicial rotándola el ángulo de la línea que mejor se ajusta. Esta inicialización resuelve el problema de quedarse atascado en mínimos locales y acelera el proceso de identificación. Por tanto, se utiliza un algoritmo de búsqueda directa en lugar de optimización genética.

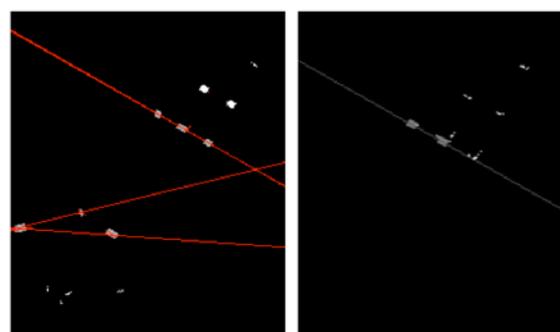


Figura 4: a) Líneas detectadas b) Línea que mejor se ajusta

La pose del pallet se acepta como buena si el máximo de puntos acoplados supera el 80 % con respecto al real. El procedimiento de inicialización puede fallar al coger objetos que no son pallets (estanterías, obstáculos, etc.) como si fueran pallets. Por ello se refuerza el algoritmo de inicialización basándose en la cámara. En vez utilizar el color que depende mucho del tipo de iluminación emplea la información geométrica.

En caso de que el candidato no supere el 80% de correlación utilizando el láser, se pregunta a la cámara y solo se acepta si los datos procesados con la cámara confirman el resultado. En la figura 4 se puede apreciar el flujo del funcionamiento:

Estos dos instrumentos se instalan en el mismo eje de forma vertical para así poder proyectar los puntos medidos por el láser en la imagen. Pero para ello filtran previamente todos los puntos del láser que no son visibles a la cámara, figura 6

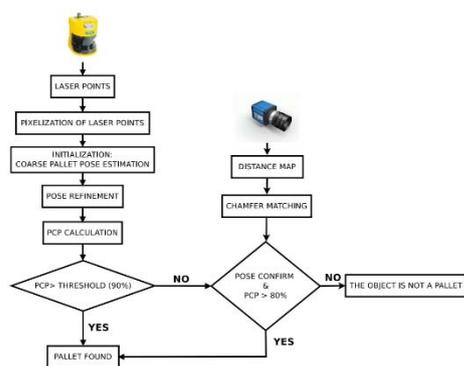


Figura 5: Diagrama de flujo del algoritmo de Baglivo[2]

En la figura 7 se puede apreciar el resultado sobre la imagen original.

Para cruzar los datos del láser con los de la cámara se utiliza la imagen de bordes extraídos, figura 8.

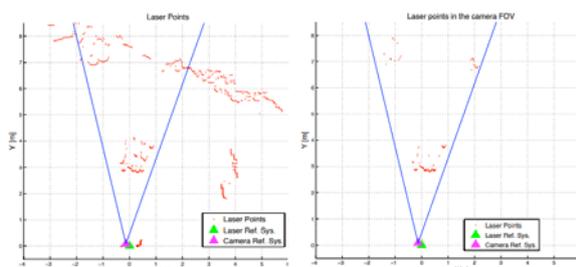


Figura 6: a) Puntos láser b) Puntos láser filtrados



Figura 7: a) Puntos láser sobre imagen original b) Puntos láser filtrados sobre imagen original

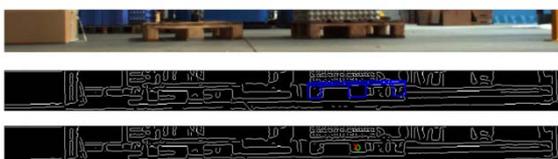


Figura 8: a) Imagen capturada de dos pallets b) Imagen de bordes con plantilla c) Solución de imagen(cruz) y solución de láser (circulo)

En este caso, el pallet se acepta ya que las dos soluciones están muy próximas. El desplazamiento entre los dos puntos se debe a fuentes de ruido, incertidumbre en los parámetros extrínsecos y ruido en la extracción de bordes.

El estudio [8] se basa en dos recursos: la intensidad de la imagen para diferenciar colores y cámaras estéreo para sacar la información 3D. Utilizan HOG (Histogram of gradient) para calcular los bordes del pallet y posteriormente normalizan ese gradiente dividiendo por un valor, los componentes del gradiente se calculan a través de una convolución con el filtro Sobel. La magnitud del gradiente tiene dos componentes x e y de la imagen. Después se pasa una ventana deslizante por toda la imagen para clasificar los candidatos posibles a pallets. Se fija el tamaño máximo del bounding box (5x20) y se elimina las zonas que no son pallet, a través de la intensidad del color. Este proceso se puede observar en la figura 9.

Después se pasa un detector de bordes cercano al límite del bounding box para mejorar la precisión de la localización. El resultado se muestra en la figura 10.

Se utiliza el LBP (Local Binary Pattern) para las características de textura, se ha demostrado que es un descriptor que es robusto antes los cambios de iluminación.

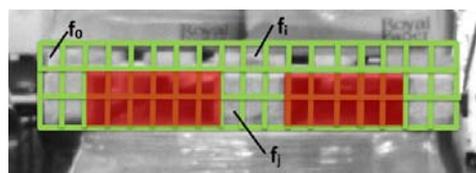


Figura 9: Detección de pallet y eliminación del hueco de los tacos

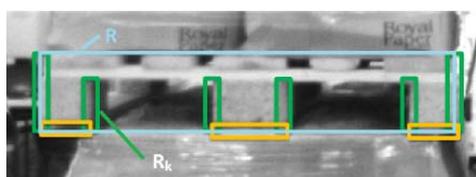


Figura 10: Detección de bordes de áreas rectangulares

En [7] se utiliza la extracción de características para la detección del pallet, posteriormente se pasa una ventana multiescala siempre con relación de aspecto de 5 que es lo que se ha obtenido experimentalmente del pallet. Es multiescala para poder detectar el pallet más cercano o más lejano. En la extracción de características lo que realmente se calcula es la magnitud del gradiente con diferentes orientaciones como se puede observar en la imagen 11. Para la clasificación, se utilizó un clasificador Adaboost.

Otro de los objetivos que tenía este proyecto era la descarga automática del pallet. Para ello era necesario detectar las zonas de descarga si están vacías o no. Se debe de detectar el paralelepípedo de descarga como el cuboide en el espacio 3D que representa el área vacía delimitada por obstáculos a la izquierda y a la derecha

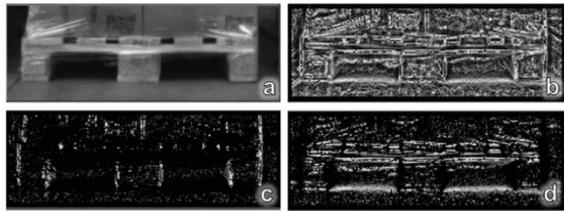


Figura 11: a) Escala de grises b) Gradiente de magnitud c) Gradiente de magnitud en la orientación 0° d) Gradiente de magnitud en la orientación 90°

Se indica un error si por cualquier razón la dimensión del paralelepípedo es menor que la dimensión del pallet. Se comienza por detectar la estructura del estante, esto se puede obtener del mapa de disparidad. Las columnas y los soportes deberían estar más o menos a la misma distancia ya que el enfoque de la cámara es perpendicular al estante. Se calcula la disparidad principal como el máximo local por encima de un cierto valor de umbral, un máximo local debe aparecer varias veces en la imagen. Una vez que se tiene la disparidad principal se filtra la imagen reteniendo los valores cercanos a la disparidad principal y el resto a cero. Posteriormente se construyen proyecciones verticales y horizontales de los elementos distintos de cero. Estas proyecciones proporcionan información necesaria para delimitar la zona libre. El resultado se muestra en la imagen 12.

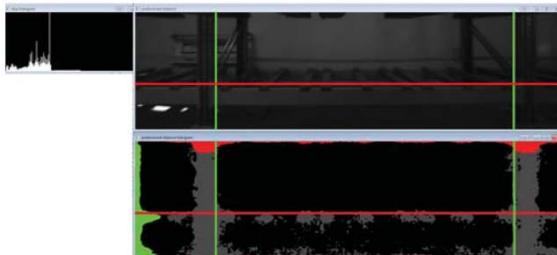


Figura 12: a) Histograma de disparidad b) Límites del cuboide de descarga dibujado en la imagen de entrada c) Imagen de disparidad con la disparidad principal resaltada

La descarga no solo se realiza en estantes si no que también se realiza en el suelo a través de marcas que delimitan la zona de descarga. En este caso se ha utilizado la transformada de Hough para detectar esas líneas filtrando con diferentes orientaciones dependiendo de la línea que queramos detectar. El resultado se muestra en la siguiente imagen 13.

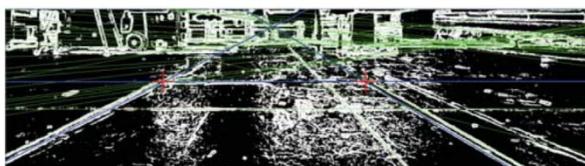


Figura 13: Líneas verdes (todas las líneas obtenidas por hough). Líneas azules (macas reales del suelo). Cruces rojas (intersección de líneas azul)

3 ALGORITMO PROPUESTO

Se ha propuesto la instalación de un sistema de visión en el AGV para dotarle de una mayor inteligencia y autonomía detectando la posición y orientación del pallet en todo momento. A continuación se explicará de manera detallada el desarrollo del estudio:

3.1 SISTEMA DE VISIÓN 3D

Se realizó un estudio comparativo para determinar qué sistema de visión 3D era el que mejor se adaptaba a los requerimientos de la aplicación. Llegando a la conclusión de que el uso de proyección de luz estructurada es más adecuado en este caso con respecto al uso de visión estereoscópica, puesto que se necesita mayor rapidez en los cálculos para que funcione en tiempo real, y además porque este sistema es robusto frente a cambios de iluminación.



Figura 13: Elección sistema de visión

La cámara de luz estructurada utilizada para la realización de pruebas fue la kinect 1.



Figura 14: Kinect 1

En la figura 15 se muestra como se comunican los diferentes componentes utilizados en este proyecto.

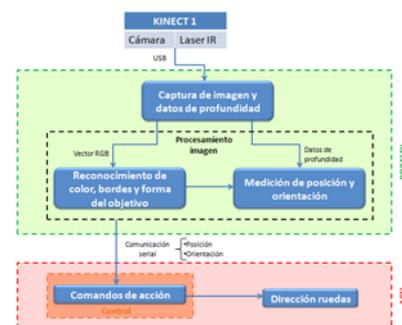


Figura 15: Comunicación entre componentes

3.2 ESTRATEGIA

Se propusieron diferentes estrategias para la detección del pallet: empleando colores y segmentación (esta opción se descartó debido a la gran variabilidad de tonos que pueden aparecer en los pallets además van cambiando a lo largo del tiempo por la naturaleza de la madera), ubicando de forma precisa códigos QR en cada esquina del pallet (esta opción se descartó puesto que los pallets en las fabricas no son fijos y no es viable la colocación de marcas artificiales), finalmente nos decantamos por realizar un filtro por tamaño.



Figura 16: Estrategia detección pallet

Como se ha comentado la selección de la estrategia radica en la necesidad de modificar lo mínimo posible el pallet original.

3.3 ALGORITMO

Para la detección de pallets se han empleado diferentes técnicas de visión por computador como por ejemplo: Transformada de Hough[6], Canny, dilatación, erosión, etc [1][5]. Se explicará por pasos de manera resumida en qué consiste el algoritmo propuesto:

1. Obtención de nube de puntos

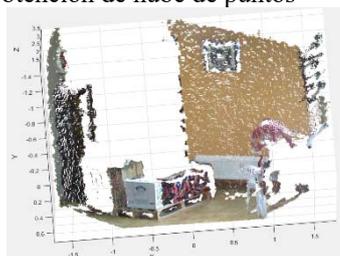


Figura 17: Nube de puntos

2. Se divide en rebanadas en Y la nube de puntos y se proyectan todos los puntos de dicha rebanada

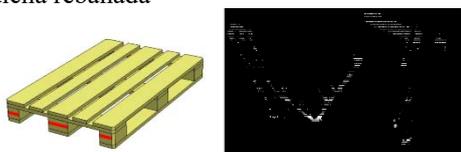


Figura 18: Rebanadas en el Eje Y

3. Proceso de cierre (Dilatación - erosión)



Figura 19: Proceso de cierre

4. Búsqueda de bounding box acorde al tamaño del pallet

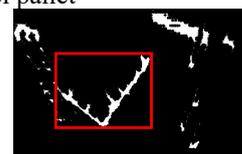


Figura 20: Bounding box

5. Cálculo de esquinas del bounding box

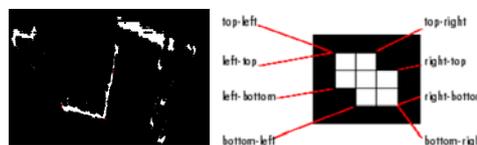


Figura 21: Cálculo de esquinas

6. Verificación de que debajo de la esquina esta el taco del pallet.



Figura 22: Tacos del pallet

7. Cálculo distancia entre esquinas. Para dar por valida la detección del pallet se tienen que dar 3 condiciones:
 - Distancia entre esquinas es acorde al tamaño real de pallet
 - Se detectan los tacos justo debajo de las esquinas.
 - Angulo recto entre los dos lados del pallet.

Una de las mayores dificultades fue tratar los casos particulares en los que solo se detectaba un lado del pallet debido a ciertas circunstancias que se pueden dar en la vida real, por ejemplo, que el pallet este con bastante carga o que el pallet este totalmente alineado con la cámara y no seamos capaces de detectar la profundidad del pallet. Para estos casos se ha preparado un árbol de decisión que nos permite detectar el pallet.

En el diagrama de flujo de la figura 23 se representa el funcionamiento básico del algoritmo.

4 RESULTADOS

Se puede avanzar que los resultados han sido satisfactorios ya que se ha obtenido casi un 92 % de acierto en la detección de pallets. La probabilidad de que se obtenga un falso positivo es de 0,01 % es un resultado bastante bajo, pero considerando que es un fallo bastante grave este error debería de evitarse por completo.

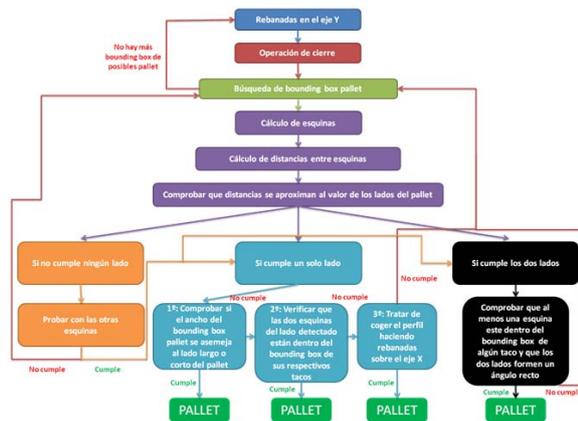


Figura 23: Diagrama de flujo del algoritmo

Se prevé que con el uso de una cámara de una precisión mayor se eliminen por completo los falsos positivos y aumente el porcentaje de éxito en la detección de pallets.

Para comprobar los resultados del algoritmo se realizaron 110 fotos adicionales, sin contar las que se hicieron para ir mejorando el algoritmo. A la hora de realizar las fotos se modificaron los siguientes parámetros de forma aleatoria:

- Variación de la altura de la cámara desde 0,1 m hasta 1,5 m medido desde el suelo. En cambio, la altura del pallet no se ha variado nunca, siempre se han hecho pruebas con el pallet en el suelo.
- Variación de la distancia del pallet a la cámara en el rango de 0,8m a 3,5m aproximadamente.
- Modificación de la orientación tanto de la cámara como del pallet.
- Diferentes entornos.
- Variación de la iluminación.
- Cambio de posición de la carga en el pallet.

La precisión de localización, como era de esperar, no ha sido tan buena como la identificación, ya que de media se produce un error de 6 cm en la posición del centro del pallet. Pero considerando que se utiliza una cámara de poca precisión y que el objetivo es

solo la detección del pallet para que posteriormente se realice la localización de manera más precisa, este resultado no es preocupante. Para la localización con precisión del pallet se recomienda emplear un láser de navegación.

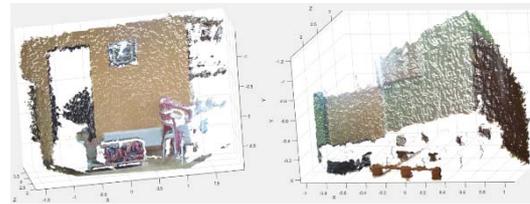


Figura 24: Variación de parámetros

5 CONCLUSIONES

Se ha realizado un estudio del estado del arte comprobando que en la actualidad no hay ningún producto comercial ya asentado que resuelva la detección de pallets.

En el mundo de la visión por computador es muy importante que la precisión de la cámara sea acorde al problema a resolver. El principal problema de la Kinect 1 es que la precisión baja considerablemente con el aumento de la distancia. Además, el rango de distancia máximo es de 4 metros, una longitud pequeña para esta aplicación.

Como alternativa a una cámara de luz estructurada, se propone instalar en la parte superior del AGV un sensor láser con un motor incorporado para poder realizar un barrido y así obtener una nube de puntos 3D mucho más precisa. Con este sistema lo que se pretende es realizar la detección y relocalización en un mismo paso con un solo sistema, para así ahorrarnos problemas de calibración y comunicación entre los dos sistemas. Se debería estudiar la relación calidad-precio de ambos sistemas y ver cuál es el sistema más adecuado.

Existen numerosas aplicaciones a desarrollar y que pueden ser muy útiles en un futuro instalando un sistema de visión en un AGV. En la actualidad la mayoría de los AGV no cuentan con un sistema de visión por lo que se podría decir que navegan "a ciegas" extrayendo muy poca información del entorno. Con la instalación de estos sistemas de visión se podría conseguir mucha más información del entorno y por lo tanto que el AGV pueda ser más autónomo e inteligente.

Se ha obtenido una solución viable, ya que se ha conseguido un porcentaje de éxito en la detección de pallets del 92 % con una cámara que no es muy precisa. En el caso de que se utilizará una cámara de mayor precisión se prevé que el porcentaje de éxito

auge considerablemente y también se prevé la reducción completa de falsos positivos. El código se ha elaborado pensando en poder reutilizarlo fácilmente, variando ciertos parámetros en el caso de que se quiera poner una cámara de mayor precisión.

English summary

PALLET DETECTION BY COMPUTER VISION TECHNIQUES

Abstract

Nowadays the AGVs know the positions of the pallets because they store their localizations. If the position of some pallet is modified by any external agent to the system, for example manned transport systems, the stored position would not match with the real one, and the AGV could not pick it. For this reason, it is needed to include optical sensors in the AGVs to identify the pallets in real time and update their trajectories. In this work a state of art on pallet detection is performed, the most suitable vision system is identified for this application, and a new algorithm based on computer vision techniques is proposed. Empirical results contrast the validity of the proposal.

Keywords: AGV, pallet, detection, algorithm

Referencias

- [1] Alegre, E., Pajares, G., & De la Escalera, A. (2016) Conceptos y Métodos en Visión por Computador, Grupo de Visión del Comité Español de Automática (CEA), España.
- [2] Baglivo, L., Biasi, N., Biral, F., Bellomo, N., Bertolazzi, E., Da Lio, M., & De Cecco, M. (2011) Autonomous pallet localization and picking for industrial forklifts: a robust range and look method, *Measurement Science and Technology*, 22(8), 085502.

- [3] Bostelman, R., Hong, T., & Chang, T. (2006, October) Visualization of pallets, In *Intelligent Robots and Computer Vision XXIV: Algorithms, Techniques, and Active Vision* (Vol. 6384, p. 638408), International Society for Optics and Photonics.
- [4] Cucchiara, R., Piccardi, M., & Prati, A. (2000, September) Focus based Feature Extraction for Pallets Recognition, In *BMVC* (pp. 1-10).
- [5] González, A., Martínez, F., Pernía, A., Alba, F., Castejón, M., Ordieres, J., & Vergara, E. (2006) *Técnicas y algoritmos básicos de visión artificial*, Universidad de la Rioja, La Rioja.
- [6] Illingworth, J., & Kittler, J. (1988) A survey of the Hough transform, *Computer vision, graphics, and image processing*, 44(1), 87-116.
- [7] Sierra, J.E. (2019) Industrial AGV holistic modelling approach, *Proceedings of 10th Congress of the Federation of European Simulation Societies, EUROSIM 2019*.
- [8] Varga, R., Costea, A., & Nedeveschi, S. (2015, September) Improved autonomous load handling with stereo cameras, In *2015 IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)* (pp. 251-256), IEEE.
- [9] Varga, R., & Nedeveschi, S. (2016, February) Robust Pallet Detection for Automated Logistics Operations, In *VISIGRAPP (4: VISAPP)* (pp. 470-477).



© 2019 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution CC BY-NC-SA 4.0 license (<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es>).